# Week 24

Shaun Stearns

9/26/2019

## Multiple Regression 1

### Lowest Weigh-in

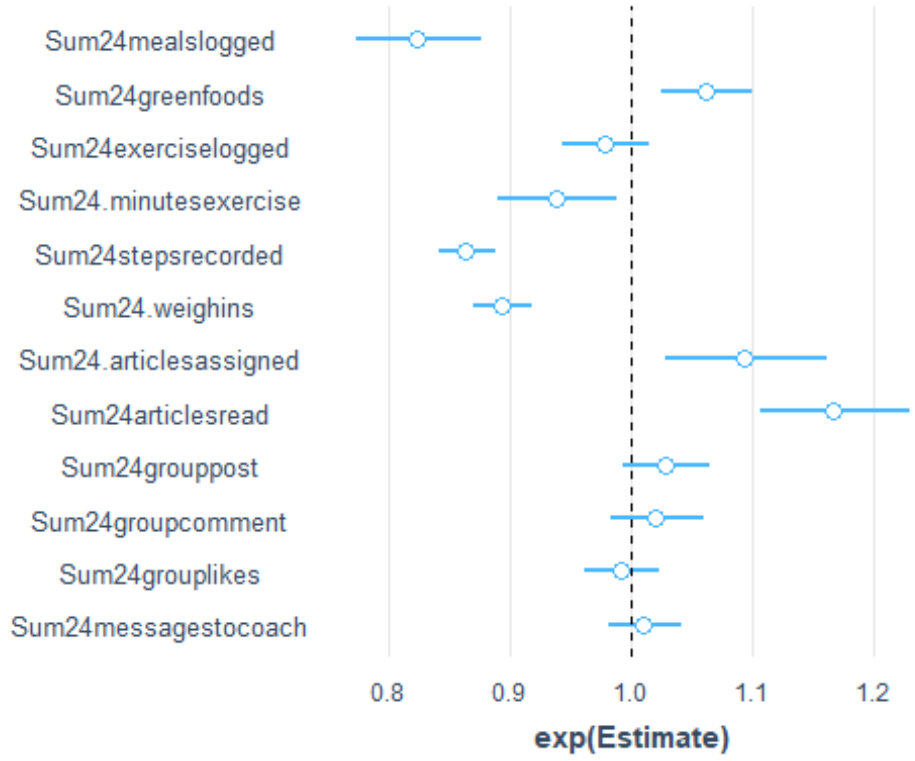
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,19,22,25,28,31,34,37,40,43,46,49,52)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(Lowestweighinkg ~., DF.t)
```

MODEL INFO:
*Observations:* 7138
*Dependent Variable:* Lowestweighinkg
*Type:* OLS linear regression

MODEL FIT:
$F(12,7125) = 30.89$, $p = 0.00$
$R^2 = 0.05$
*Adj.* $R^2 = 0.05$

*Standard errors: OLS*

| | Est. | 2.5% | 97.5% | t val. | p | partial.r | part.r |
|---|---|---|---|---|---|---|---|
| (Intercept) | 0.00 | -0.02 | 0.02 | 0.00 | 1.00 | | |
| Sum24mealslogged | -0.19 | -0.26 | -0.13 | -6.04 | 0.00 | -0.07 | -0.07 |
| Sum24greenfoods | 0.06 | 0.03 | 0.09 | 3.89 | 0.00 | 0.05 | 0.04 |
| Sum24exerciselogged | -0.02 | -0.06 | 0.01 | -1.21 | 0.23 | -0.01 | -0.01 |
| Sum24.minutesexercise | -0.06 | -0.10 | -0.03 | -3.94 | 0.00 | -0.05 | -0.05 |
| Sum24stepsrecorded | -0.15 | -0.18 | -0.12 | -9.74 | 0.00 | -0.11 | -0.11 |
| Sum24.weighins | -0.11 | -0.14 | -0.08 | -7.93 | 0.00 | -0.09 | -0.09 |
| Sum24.articlesassigned | 0.09 | 0.03 | 0.15 | 2.96 | 0.00 | 0.04 | 0.03 |
| Sum24articlesread | 0.15 | 0.10 | 0.21 | 5.61 | 0.00 | 0.07 | 0.06 |
| Sum24grouppost | 0.03 | -0.00 | 0.06 | 1.70 | 0.09 | 0.02 | 0.02 |
| Sum24groupcomment | 0.02 | -0.02 | 0.06 | 1.02 | 0.31 | 0.01 | 0.01 |
| Sum24grouplikes | -0.01 | -0.04 | 0.03 | -0.47 | 0.64 | -0.01 | -0.01 |
| Sum24messagestocoach | 0.01 | -0.02 | 0.04 | 0.67 | 0.50 | 0.01 | 0.01 |



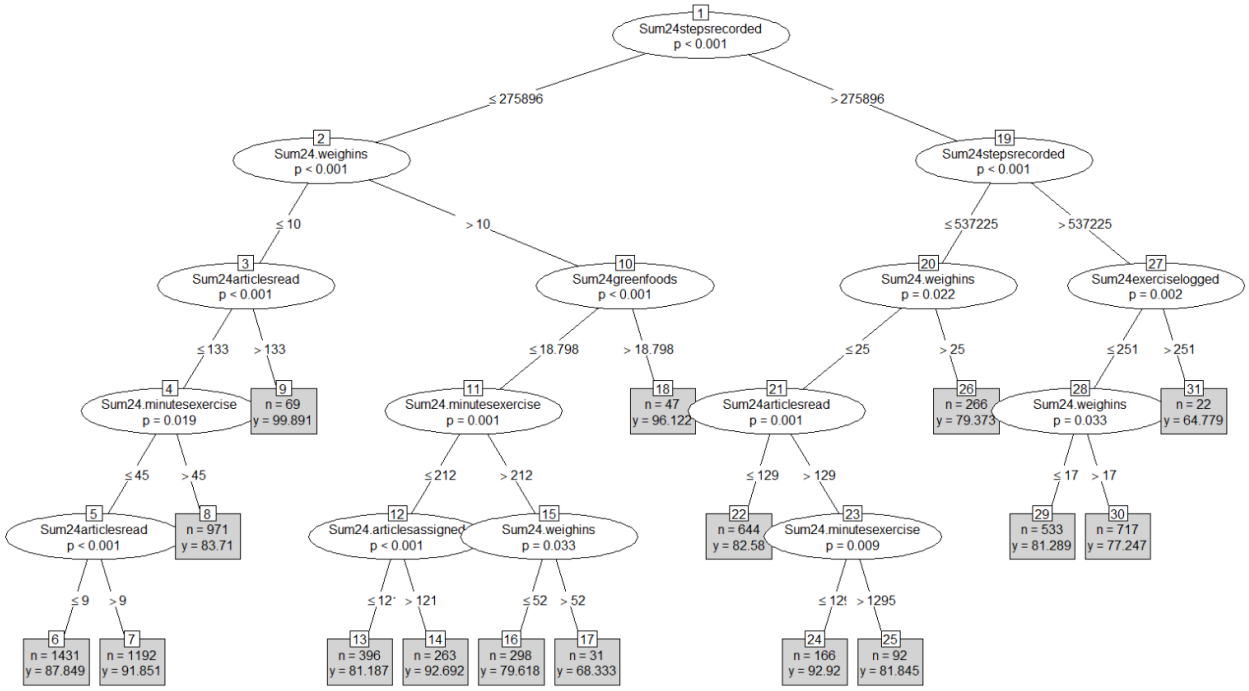

## Regression Tree 1

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,19,22,25,28,31,34,37,40,43,46,49,52)]
model <- train(
  Lowestweighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD  RsquaredSD
## 1         0.95 20.26493 0.03390884 15.65617 0.5257023 0.008378094
##       MAESD
## 1 0.3041787
```

## Tree Model

```
plot(model$finalModel, type = "simple")
```




## User Engagement and Lowest Weigh-in based on 16 Terminal Nodes (Left to Right)

| | |
|---|---|
| **1** | Lowest Weigh-in avg 87.85 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins ≤ 10, WK 24 Articles Read ≤ 133, WK 24 Min of Exercise ≤ 45, WK 24 Articles Read ≤ 9** |
| **2** | Lowest Weigh-in avg 91.85 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins ≤ 10, WK 24 Articles Read ≤ 133, WK 24 Min of Exercise ≤ 45, WK 24 Articles Read > 9** |
| **3** | Lowest Weigh-in avg 83.71 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins ≤ 10, WK 24 Articles Read ≤ 133, WK 24 Min of Exercise > 45** |
| **4 High** | Lowest Weigh-in avg 99.89 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins ≤ 10, WK 24 Articles Read > 133** |
| **5** | Lowest Weigh-in avg 81.19 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins > 10, WK 24 Green Foods ≤ 18.80, WK 24 Min of Exercise ≤ 212, WK 24 Articles Assigned ≤ 121** |
| **6** | Lowest Weigh-in avg 92.69 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins > 10, WK 24 Green Foods ≤ 18.80, WK 24 Min of Exercise ≤ 212, WK 24 Articles Assigned > 121** |
| **7** | Lowest Weigh-in avg 79.62 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins > 10, WK 24 Green Foods ≤ 18.80, WK 24 Min of Exercise > 212, WK 24 Weigh-ins ≤ 52** |
| **8 2nd Low** | Lowest Weigh-in avg 68.33 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins > 10, WK 24 Green Foods ≤ 18.80, WK 24 Min of Exercise > 212, WK 24 Weigh-ins > 52** |
| **9** | Lowest Weigh-in avg 96.12 kg = **WK 24 Steps Recorded ≤ 275896, WK 24 Weigh-ins > 10, WK 24 Green Foods > 18.80** |
| **10** | Lowest Weigh-in avg 82.58 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded ≤ 537225, WK 24 Weigh-ins ≤ 25, WK 24 Articles Read ≤ 129** |
| **11** | Lowest Weigh-in avg 92.92 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded ≤ 537225, WK 24 Weigh-ins ≤ 25, WK 24 Articles Read > 129, WK 24 Min of Exercise ≤ 1295** |
| **12** | Lowest Weigh-in avg 81.85 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded ≤ 537225, WK 24 Weigh-ins ≤ 25, WK 24 Articles Read > 129, WK 24 Min of Exercise > 1295** |

| | |
|---|---|
| **13** | Lowest Weigh-in avg 79.37 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded ≤ 537225, WK 24 Weigh-ins > 25** |
| **14** | Lowest Weigh-in avg 81.29 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded > 537225, WK 24 Exercise Logged ≤ 251, WK 24 Weigh-ins ≤ 17** |
| **15** | Lowest Weigh-in avg 77.25 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded > 537225, WK 24 Exercise Logged ≤ 251, WK 24 Weigh-ins > 17** |
| **16 Low** | Lowest Weigh-in avg 64.78 kg = **WK 24 Steps Recorded > 275896, WK 24 Steps Recorded > 537225, WK 24 Exercise Logged > 251** |

## Random Forest 1

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,19,22,25,28,31,34,37,40,43,46,49,52)]
rf <- randomForest(Lowestweighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

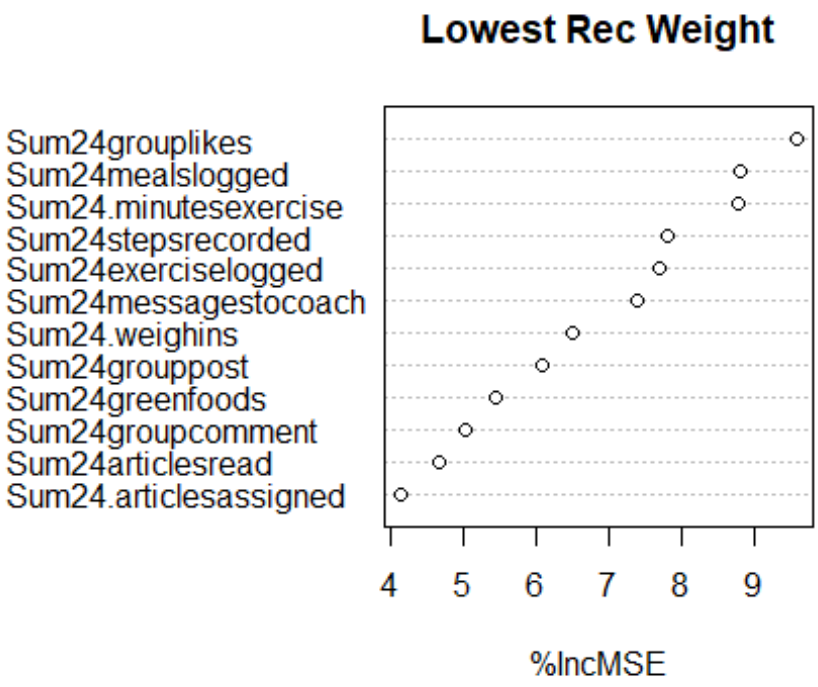
## Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.1580012
##   Random R-square:  -0.2664959
##   Random R-square variance:  0.0001851399
```

## Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Lowest Rec Weight")
```




## Multiple Regression 2

### Difference Between First Weigh-in and Lowest Weigh-in

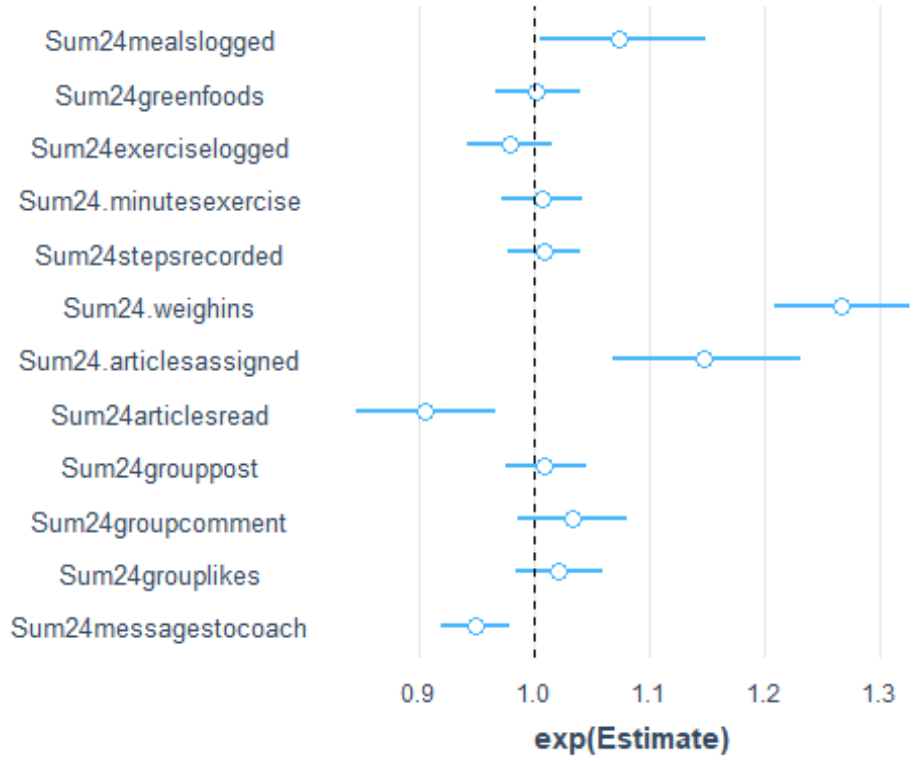
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,19,22,25,28,31,34,37,40,43,46,49,52)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(AbsDiffFirstWeighinkg ~., DF.t)
```

MODEL INFO:
*Observations:* 7138
*Dependent Variable:* AbsDiffFirstWeighinkg
*Type:* OLS linear regression

MODEL FIT:
$F(12,7125) = 64.23$, $p = 0.00$
$R^2 = 0.10$
*Adj.* $R^2 = 0.10$

*Standard errors: OLS*

| | Est. | 2.5% | 97.5% | t val. | p | partial.r | part.r |
|---|---|---|---|---|---|---|---|
| (Intercept) | -0.00 | -0.02 | 0.02 | -0.00 | 1.00 | | |
| Sum24mealslogged | 0.07 | 0.01 | 0.13 | 2.27 | 0.02 | 0.03 | 0.03 |
| Sum24greenfoods | 0.00 | -0.03 | 0.03 | 0.15 | 0.88 | 0.00 | 0.00 |
| Sum24exerciselogged | -0.02 | -0.06 | 0.01 | -1.29 | 0.20 | -0.02 | -0.01 |
| Sum24.minutesexercise | 0.01 | -0.03 | 0.04 | 0.38 | 0.70 | 0.00 | 0.00 |
| Sum24stepsrecorded | 0.01 | -0.02 | 0.04 | 0.56 | 0.58 | 0.01 | 0.01 |
| Sum24.weighins | 0.24 | 0.21 | 0.26 | 17.07 | 0.00 | 0.20 | 0.19 |
| Sum24.articlesassigned | 0.14 | 0.08 | 0.19 | 4.68 | 0.00 | 0.06 | 0.05 |
| Sum24articlesread | -0.10 | -0.15 | -0.05 | -3.74 | 0.00 | -0.04 | -0.04 |
| Sum24grouppost | 0.01 | -0.02 | 0.04 | 0.56 | 0.58 | 0.01 | 0.01 |
| Sum24groupcomment | 0.03 | -0.01 | 0.07 | 1.64 | 0.10 | 0.02 | 0.02 |
| Sum24grouplikes | 0.02 | -0.01 | 0.05 | 1.17 | 0.24 | 0.01 | 0.01 |
| Sum24messagestocoach | -0.05 | -0.08 | -0.02 | -3.39 | 0.00 | -0.04 | -0.04 |




## Regression Tree 2

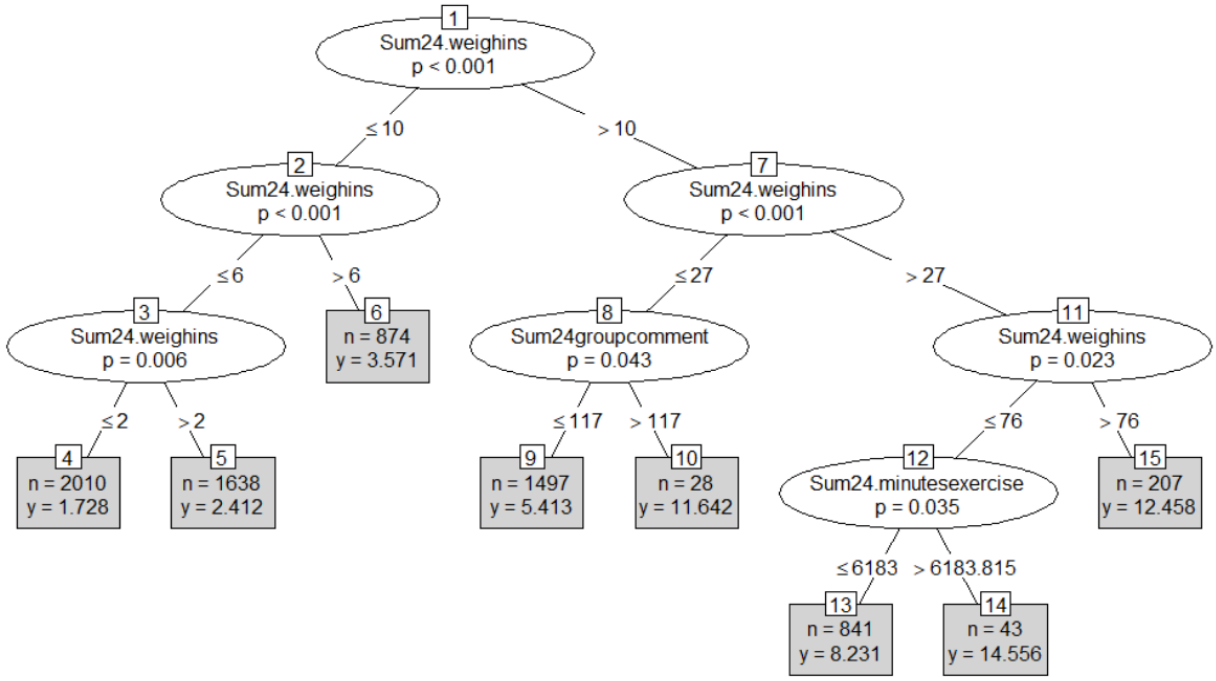
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header = TRUE)
```

```
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,19,22,25,28,31,34,37,40,43,46,49,52)]
model <- train(
  AbsDiffFirstWeighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE  Rsquared      MAE    RMSESD RsquaredSD     MAESD
## 1         0.95 7.697325 0.1018695 3.824067 0.7949311 0.02815181 0.2981755
```

## Tree PLot

```
plot(model$finalModel, type = "simple")
```




## User Engagement and Difference Between First Weigh-in and Lowest Weigh-in based on 8 Terminal Nodes (Left to Right)

| | |
|---|---|
| **1 High** | Diff 1st Weight Lost avg 1.73 kg = **WK 24 Weigh-ins ≤ 2** |
| **2** | Diff 1st Weight Lost avg 2.41 kg = **WK 24 Weigh-ins > 2 & ≤ 6** |
| **3** | Diff 1st Weight Lost avg 3.57 kg = **WK 24 Weigh-ins > 6 & ≤ 10** |
| **4** | Diff 1st Weight Lost avg 5.41 kg = **WK 24 Weigh-ins > 10, WK 24 Weigh-ins ≤ 27, WK 24 Group Comments ≤ 117** |
| **5** | Diff 1st Weight Lost avg 11.64 kg = **WK 24 Weigh-ins > 10, WK 24 Weigh-ins ≤ 27, WK 24 Group Comments > 117** |

| | |
|---|---|
| **6** | Diff 1st Weight Lost avg 8.23 kg = **WK 24 Weigh-ins > 10, WK 24 Weigh-ins > 27, WK 24 Weigh-ins ≤ 76, WK 24 Min of Exercise ≤ 6183.82** |
| **7 Low** | Diff 1st Weight Lost avg 14.56 kg = **WK 24 Weigh-ins > 10, WK 24 Weigh-ins > 27, WK 24 Weigh-ins ≤ 76, WK 24 Min of Exercise > 6183.82** |
| **8 2nd Low** | Diff 1st Weight Lost avg 12.49 kg = **WK 24 Weigh-ins > 76** |

## Random Forest 2

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,19,22,25,28,31,34,37,40,43,46,49,52)]
rf <- randomForest(AbsDiffFirstWeighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

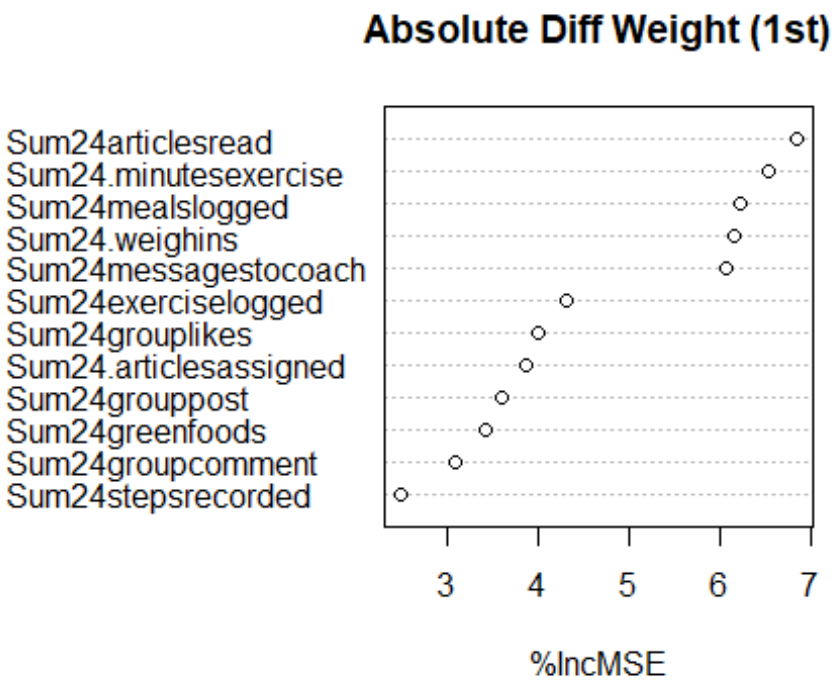
## Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.09520746
##   Random R-square:  -0.2804648
##   Random R-square variance:  0.0004341879
```

## Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Absolute Diff Weight (1st)")
```




## Multiple Regression 3

### Differnece Between Initial Weigh-in and Lowest Weigh-in

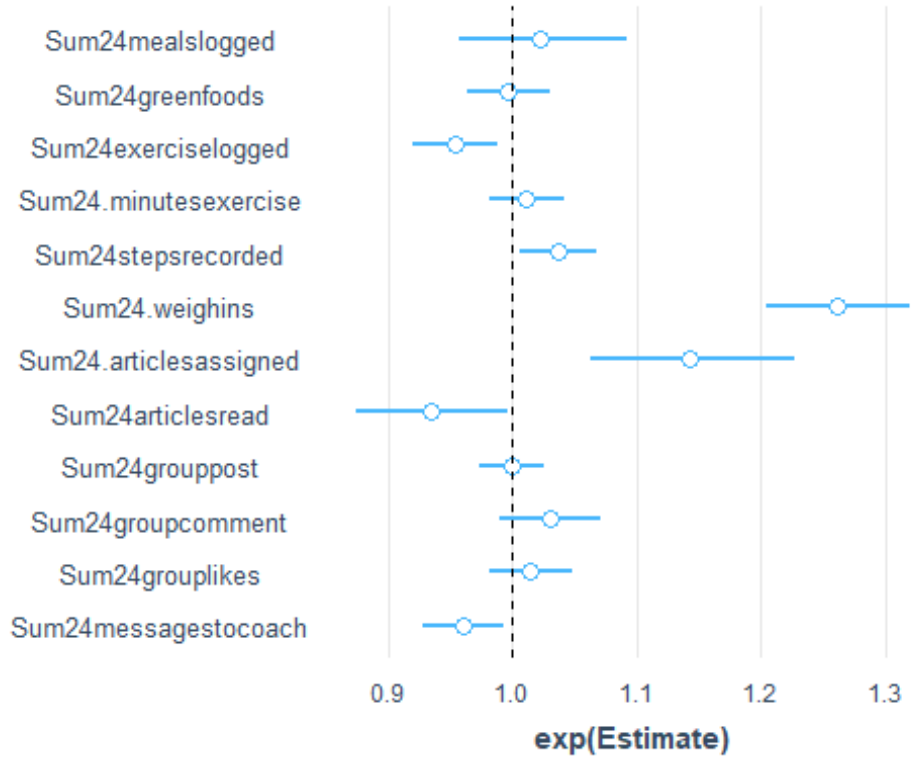
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,19,22,25,28,31,34,37,40,43,46,49,52)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(AbsDiffInitWeighinkg ~., DF.t)
```

MODEL INFO:
*Observations:* 7138
*Dependent Variable:* AbsDiffInitWeighinkg
*Type:* OLS linear regression

MODEL FIT:
$F(12,7125) = 54.80$, $p = 0.00$
$R^2 = 0.08$
*Adj.* $R^2 = 0.08$

*Standard errors: OLS*

| | Est. | 2.5% | 97.5% | t val. | p | partial.r | part.r |
|---|---|---|---|---|---|---|---|
| (Intercept) | -0.00 | -0.02 | 0.02 | -0.00 | 1.00 | | |
| Sum24mealslogged | 0.02 | -0.04 | 0.08 | 0.68 | 0.50 | 0.01 | 0.01 |
| Sum24greenfoods | -0.00 | -0.03 | 0.03 | -0.27 | 0.79 | -0.00 | -0.00 |
| Sum24exerciselogged | -0.05 | -0.08 | -0.01 | -2.74 | 0.01 | -0.03 | -0.03 |
| Sum24.minutesexercise | 0.01 | -0.02 | 0.04 | 0.68 | 0.49 | 0.01 | 0.01 |
| Sum24stepsrecorded | 0.04 | 0.01 | 0.06 | 2.40 | 0.02 | 0.03 | 0.03 |
| Sum24.weighins | 0.23 | 0.20 | 0.26 | 16.70 | 0.00 | 0.19 | 0.19 |
| Sum24.articlesassigned | 0.13 | 0.07 | 0.19 | 4.51 | 0.00 | 0.05 | 0.05 |
| Sum24articlesread | -0.07 | -0.12 | -0.02 | -2.55 | 0.01 | -0.03 | -0.03 |
| Sum24grouppost | -0.00 | -0.03 | 0.03 | -0.08 | 0.94 | -0.00 | -0.00 |
| Sum24groupcomment | 0.03 | -0.01 | 0.07 | 1.51 | 0.13 | 0.02 | 0.02 |
| Sum24grouplikes | 0.01 | -0.02 | 0.05 | 0.80 | 0.43 | 0.01 | 0.01 |
| Sum24messagestocoach | -0.04 | -0.07 | -0.01 | -2.59 | 0.01 | -0.03 | -0.03 |




## Regression Tree 3

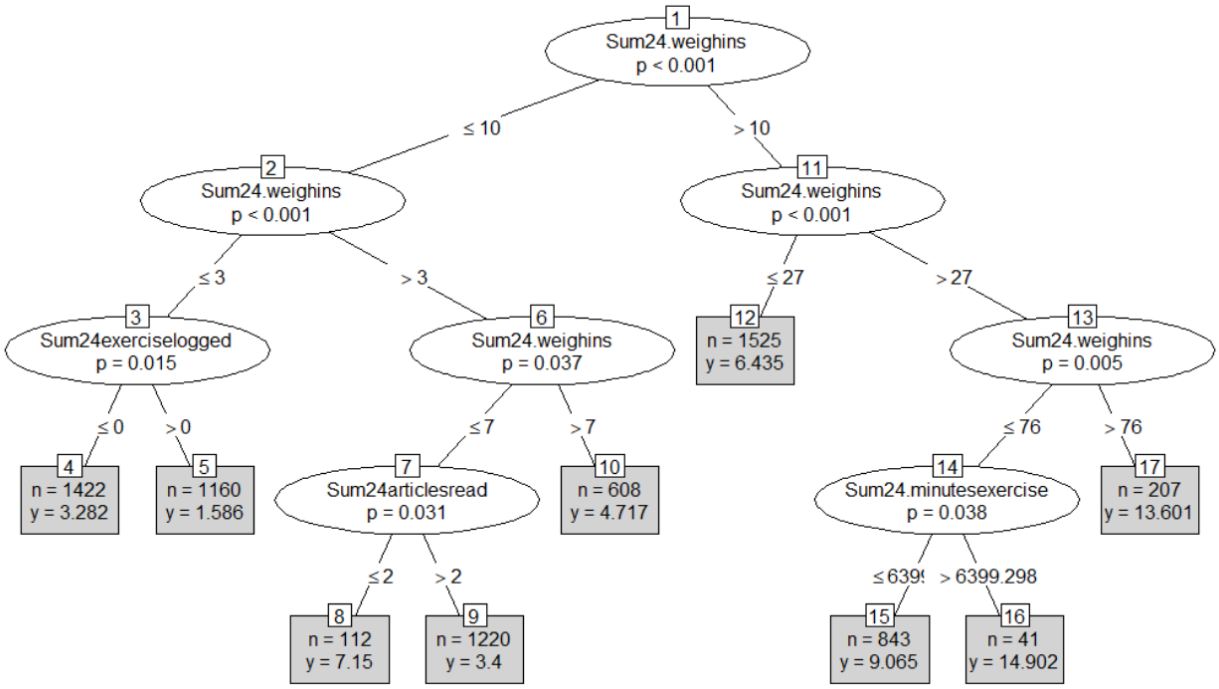
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
```

```
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,19,22,25,28,31,34,37,40,43,46,49,52)]
model <- train(
  AbsDiffInitWeighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE   RMSESD RsquaredSD     MAESD
## 1         0.95 8.331437 0.09553066 4.310869 1.110579 0.02254139 0.2491405
```

## Tree Model

```
plot(model$finalModel, type = "simple")
```




## User Engagement and Difference Between Initial Weigh-in and Lowest Weigh-in based on 9 Terminal Nodes (Left to Right)

| | |
|---|---|
| **1** | Diff Initial Weight Lost avg 3.28 kg = **WK 24 Weigh-ins ≤ 3, WK 24 Exercise Logged ≤ 0** |
| **2 High** | Diff Initial Weight Lost avg 1.59 kg = **WK 24 Weigh-ins ≤ 3, WK 24 Exercise Logged > 0** |
| **3** | Diff Initial Weight Lost avg 7.15 kg = **WK 24 Weigh-ins ≤ 10, WK 24 Weigh-ins > 3, WK 24 Weigh-ins ≤ 7, WK 24 Articles Read ≤ 2** |
| **4** | Diff Initial Weight Lost avg 3.4 kg = **WK 24 Weigh-ins ≤ 10, WK 24 Weigh-ins > 3, WK 24 Weigh-ins ≤ 7, WK 24 Articles Read > 2** |

| | |
|---|---|
| **5** | Diff Initial Weight Lost avg 4.72 kg = **WK 24 Weigh-ins ≤ 10, WK 24 Weigh-ins > 3, WK 24 Weigh-ins > 7** |
| **6** | Diff Initial Weight Lost avg 6.44 kg = **WK 24 Weigh-ins > 10, WK 24 Weigh-ins ≤ 27** |
| **7** | Diff Initial Weight Lost avg 9.09 kg = **WK 24 Weigh-ins > 27, WK 24 Weigh-ins ≤ 76, WK 24 Minutes of Exercise ≤ 6399.30** |
| **8 Low** | Diff Initial Weight Lost avg 14.90 kg = **WK 24 Weigh-ins > 27, WK 24 Weigh-ins ≤ 76, WK 24 Minutes of Exercise > 6399.30** |
| **9 2nd Low** | Diff Initial Weight Lost avg 13.60 kg = **WK 24 Weigh-ins > 76** |

## Random Forest 3

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,19,22,25,28,31,34,37,40,43,46,49,52)]
rf <- randomForest(AbsDiffInitWeighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

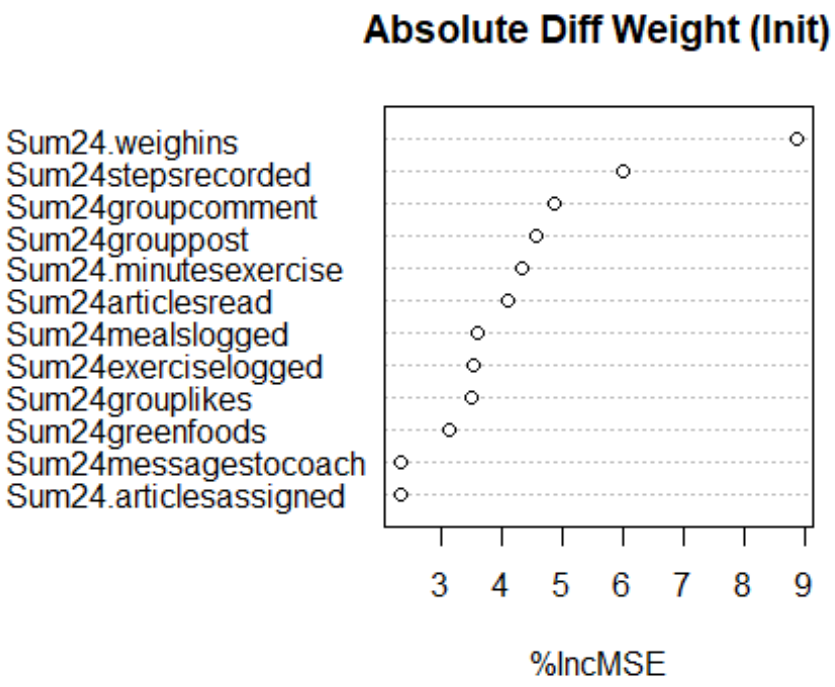
## Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.1148524
##   Random R-square:  -0.2823863
##   Random R-square variance:  0.000361325
```

## Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Absolute Diff Weight (Init)")
```




## Multiple Regression 4

### Curriculum Week (or Length of Time with DF?)

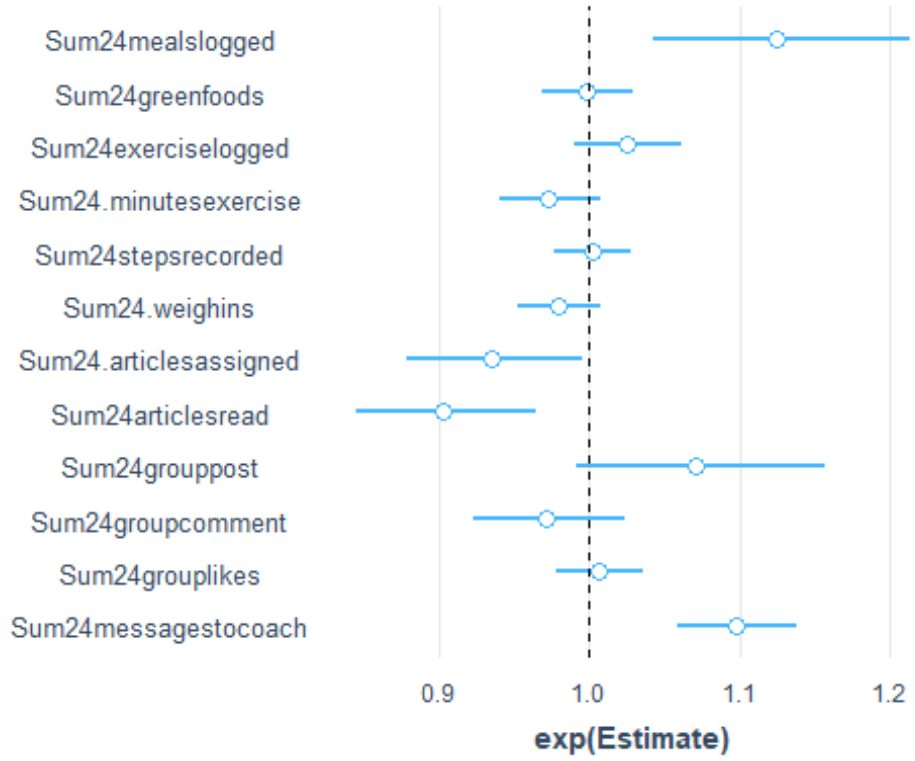
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,19,22,25,28,31,34,37,40,43,46,49,52)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(CurriculumWeek ~., DF.t)
```

MODEL INFO:
*Observations:* 7138
*Dependent Variable:* CurriculumWeek
*Type:* OLS linear regression

MODEL FIT:
$F(12,7125) = 7.00$, $p = 0.00$
$R^2 = 0.01$
*Adj.* $R^2 = 0.01$

*Standard errors: OLS*

| | Est. | 2.5% | 97.5% | t val. | p | partial.r | part.r |
|---|---|---|---|---|---|---|---|
| (Intercept) | -0.00 | -0.02 | 0.02 | -0.00 | 1.00 | | |
| Sum24mealslogged | 0.12 | 0.05 | 0.18 | 3.58 | 0.00 | 0.04 | 0.04 |
| Sum24greenfoods | -0.00 | -0.03 | 0.03 | -0.11 | 0.91 | -0.00 | -0.00 |
| Sum24exerciselogged | 0.03 | -0.01 | 0.06 | 1.38 | 0.17 | 0.02 | 0.02 |
| Sum24.minutesexercise | -0.03 | -0.06 | 0.01 | -1.64 | 0.10 | -0.02 | -0.02 |
| Sum24stepsrecorded | 0.00 | -0.03 | 0.03 | 0.09 | 0.93 | 0.00 | 0.00 |
| Sum24.weighins | -0.02 | -0.05 | 0.01 | -1.44 | 0.15 | -0.02 | -0.02 |
| Sum24.articlesassigned | -0.07 | -0.13 | -0.01 | -2.19 | 0.03 | -0.03 | -0.03 |
| Sum24articlesread | -0.10 | -0.16 | -0.05 | -3.65 | 0.00 | -0.04 | -0.04 |
| Sum24grouppost | 0.07 | 0.03 | 0.10 | 4.00 | 0.00 | 0.05 | 0.05 |
| Sum24groupcomment | -0.03 | -0.07 | 0.01 | -1.43 | 0.15 | -0.02 | -0.02 |
| Sum24grouplikes | 0.01 | -0.03 | 0.04 | 0.33 | 0.74 | 0.00 | 0.00 |
| Sum24messagestocoach | 0.09 | 0.06 | 0.12 | 5.67 | 0.00 | 0.07 | 0.07 |




## Regression Tree 4

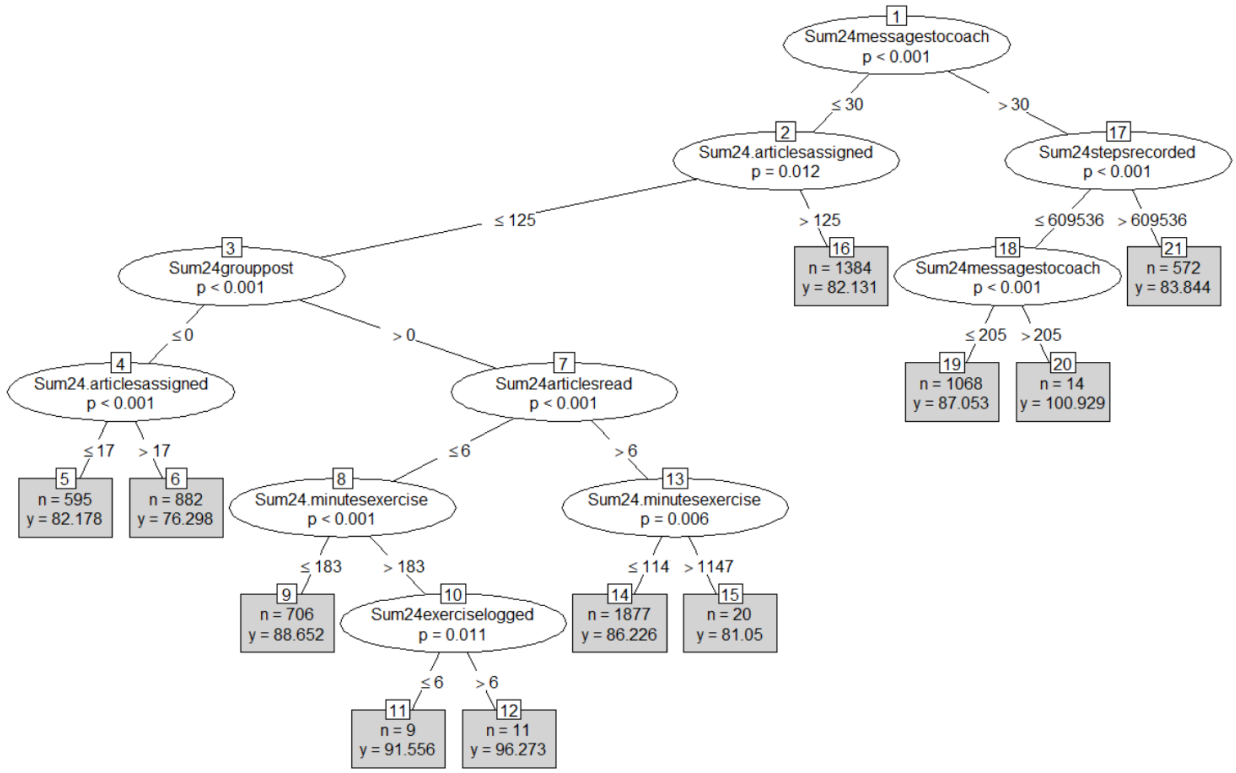
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header = TRUE)
```

```
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,19,22,25,28,31,34,37,40,43,46,49,52)]
model <- train(
  CurriculumWeek ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD RsquaredSD     MAESD
## 1         0.95 14.43246 0.04883113 8.589131 0.8636388 0.02316516 0.3559274
```

## Tree Plot

```
plot(model$finalModel, type = "simple")
```




## User Engagement and Curriculum Week based on 11 Terminal Nodes (Left to Right)

| | |
|---|---|
| **1** | Curriculum Week avg 82.18 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts ≤ 0, WK 16 Articles Assigned ≤ 17** |
| **2 Low** | Curriculum Week avg 76.30 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts ≤ 0, WK 24 Articles Assigned > 17** |

| | |
|---|---|
| **3** | Curriculum Week avg 88.65 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts > 0, WK 24 Articles Read ≤ 6, WK 24 Minutes Exercise ≤ 109, WK 24 Articles Read ≤ 6, WK 24 Min of Exercise ≤ 183** |
| **4** | Curriculum Week avg 91.56 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts > 0, WK 24 Articles Read ≤ 6, WK 24 Minutes Exercise ≤ 109, WK 24 Articles Read ≤ 6, WK 24 Min of Exercise > 183, WK 24 Exercises Logged ≤ 6** |
| **5** | Curriculum Week avg 96.27 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts > 0, WK 24 Articles Read ≤ 6, WK 24 Minutes Exercise ≤ 109, WK 24 Articles Read ≤ 6, WK 24 Min of Exercise > 183, WK 24 Exercises Logged > 6** |
| **6** | Curriculum Week avg 86.23 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts > 0, WK 24 Articles Read > 6, WK 24 Min of Exercise ≤ 114** |
| **7** | Curriculum Week avg 81.05 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned ≤ 125, WK 24 Group Posts > 0, WK 24 Articles Read > 6, WK 24 Min of Exercise > 114** |
| **8** | Curriculum Week avg 82.13 = **WK 24 Messages to Coach ≤ 30, WK 24 Articles Assigned > 125** |
| **9** | Curriculum Week avg 87.05 = **WK 24 Messages to Coach ≤ 30, WK 24 Steps Recorded ≤ 609536, WK 24 Messages to Coach ≤ 205** |
| **10 High** | Curriculum Week avg 100.93 = **WK 24 Messages to Coach ≤ 30, WK 24 Steps Recorded ≤ 609536, WK 24 Messages to Coach > 205** |
| **11** | Curriculum Week avg 83.84 = **WK 24 Messages to Coach ≤ 30, WK 24 Steps Recorded ≤ 609536** |

## Random Forest 4

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,19,22,25,28,31,34,37,40,43,46,49,52)]
rf <- randomForest(CurriculumWeek ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

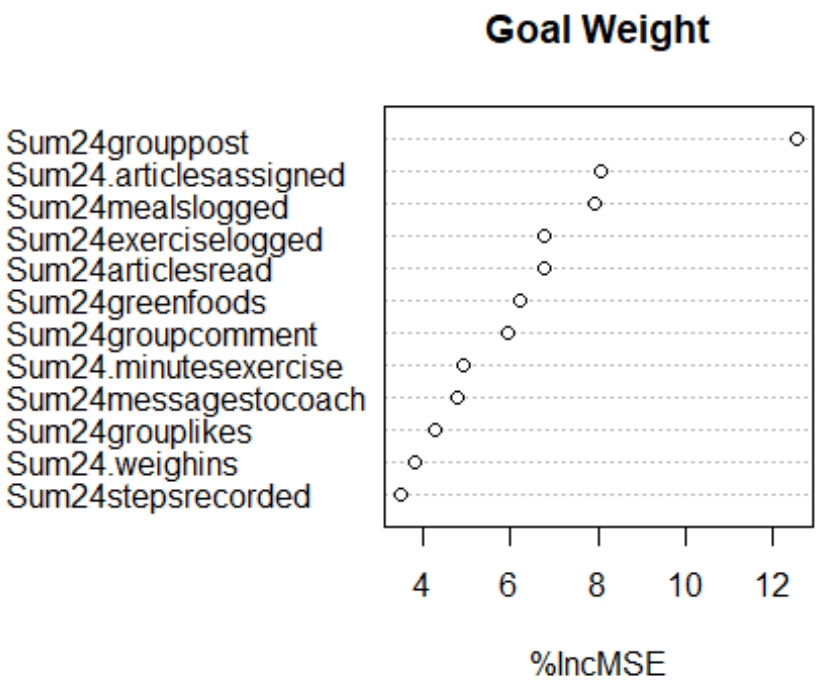
## Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.1215438
##   Random R-square:  -0.2750834
##   Random R-square variance:  0.0002523352
```

## Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Goal Weight")
```




## Multiple Regression 5

### Differnce in BMI

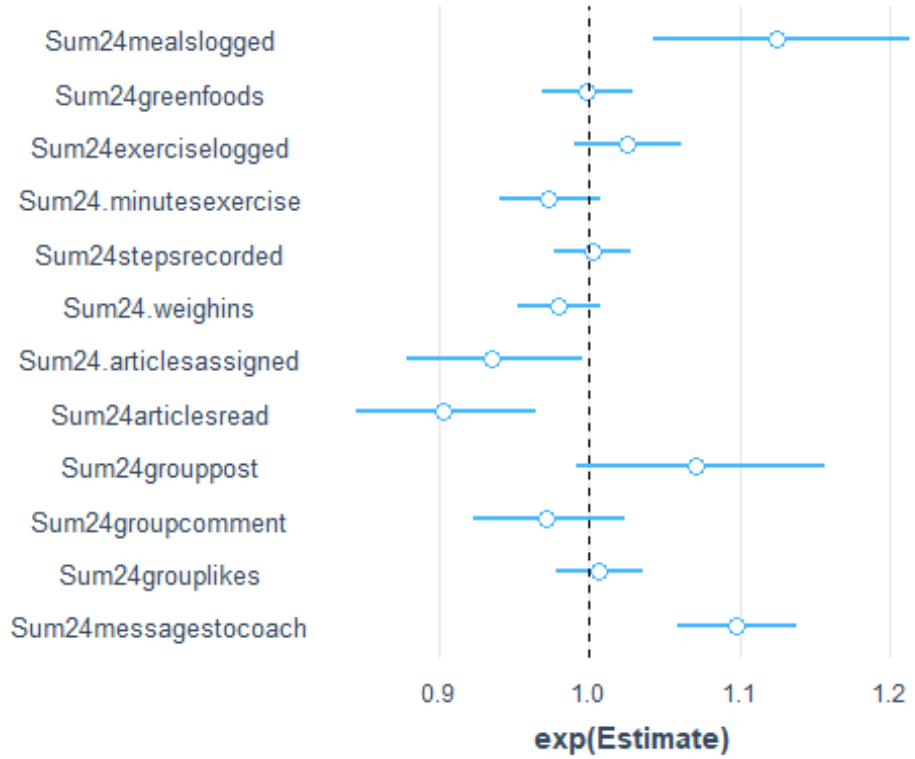
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,19,22,25,28,31,34,37,40,43,46,49,52)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(CurriculumWeek ~., DF.t)
```

MODEL INFO:
*Observations:* 7138
*Dependent Variable:* CurriculumWeek
*Type:* OLS linear regression

MODEL FIT:
$F(12,7125) = 7.00$, $p = 0.00$
$R^2 = 0.01$
*Adj.* $R^2 = 0.01$

*Standard errors: OLS*

| | Est. | 2.5% | 97.5% | t val. | p | partial.r | part.r |
|---|---|---|---|---|---|---|---|
| (Intercept) | -0.00 | -0.02 | 0.02 | -0.00 | 1.00 | | |
| Sum24mealslogged | 0.12 | 0.05 | 0.18 | 3.58 | 0.00 | 0.04 | 0.04 |
| Sum24greenfoods | -0.00 | -0.03 | 0.03 | -0.11 | 0.91 | -0.00 | -0.00 |
| Sum24exerciselogged | 0.03 | -0.01 | 0.06 | 1.38 | 0.17 | 0.02 | 0.02 |
| Sum24.minutesexercise | -0.03 | -0.06 | 0.01 | -1.64 | 0.10 | -0.02 | -0.02 |
| Sum24stepsrecorded | 0.00 | -0.03 | 0.03 | 0.09 | 0.93 | 0.00 | 0.00 |
| Sum24.weighins | -0.02 | -0.05 | 0.01 | -1.44 | 0.15 | -0.02 | -0.02 |
| Sum24.articlesassigned | -0.07 | -0.13 | -0.01 | -2.19 | 0.03 | -0.03 | -0.03 |
| Sum24articlesread | -0.10 | -0.16 | -0.05 | -3.65 | 0.00 | -0.04 | -0.04 |
| Sum24grouppost | 0.07 | 0.03 | 0.10 | 4.00 | 0.00 | 0.05 | 0.05 |
| Sum24groupcomment | -0.03 | -0.07 | 0.01 | -1.43 | 0.15 | -0.02 | -0.02 |
| Sum24grouplikes | 0.01 | -0.03 | 0.04 | 0.33 | 0.74 | 0.00 | 0.00 |
| Sum24messagestocoach | 0.09 | 0.06 | 0.12 | 5.67 | 0.00 | 0.07 | 0.07 |




## Regression Tree 5

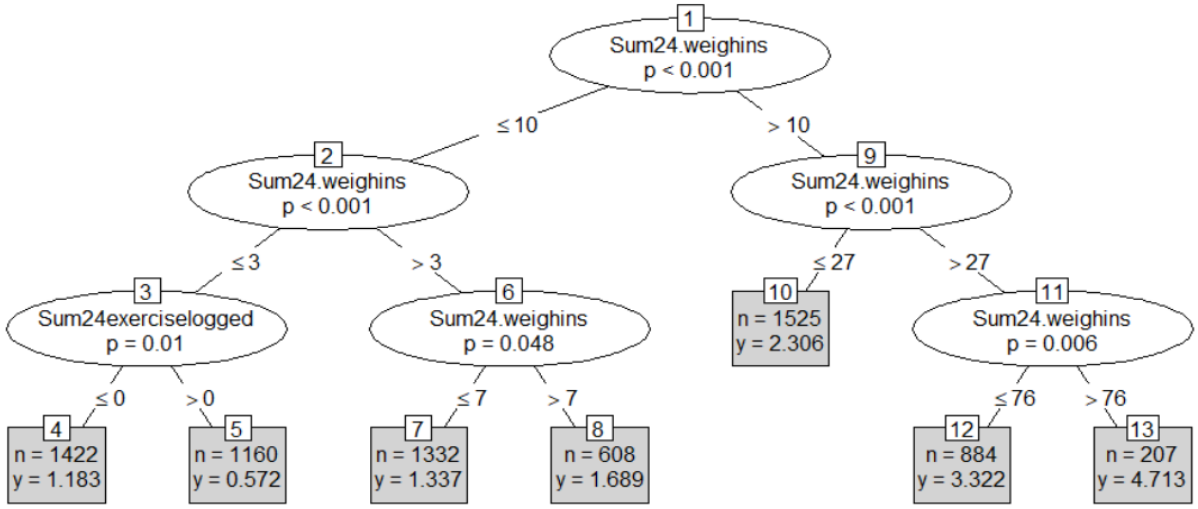
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
```

```
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,19,22,25,28,31,34,37,40,43,46,49,52)]
model <- train(
  BMIDifference ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD RsquaredSD
## 1         0.95 2.931273 0.09638345 1.528774 0.3476214 0.02531386
##        MAESD
## 1 0.07978608
```

## Tree Model

```
plot(model$finalModel, type = "simple")
```




## User Engagement and Change in BMI for 7 Terminal Nodes (Left to Right)

| | |
|---|---|
| **1** | Diff BMI avg 1.18 = **WK 24 Weigh-ins ≤ 3, WK 24 Exercise Logged ≤ 0** |
| **2 Low** | Diff BMI avg 0.57 = **WK 24 Weigh-ins ≤ 3, WK 24 Exercise Logged > 0** |
| **3** | Diff BMI avg 1.34 = **WK 24 Weigh-ins ≤ 7** |
| **4** | Diff BMI avg 1.69 = **WK 24 Weigh-ins > 7 & ≤ 10** |
| **5** | Diff BMI avg 2.31 = **WK 24 Weigh-ins > 10 & ≤ 27** |
| **6** | Diff BMI avg 3.32 = **WK 24 Weigh-ins > 27 & ≤ 76** |
| **7 High** | Diff BMI avg 4.71 = **WK 24 Weigh-ins > 76** |

## Random Forest 5

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,19,22,25,28,31,34,37,40,43,46,49,52)]
rf <- randomForest(BMIDifference ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

## Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.09987215
##   Random R-square:  -0.2756045
##   Random R-square variance:  0.0003885378
```

## Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "BMI Diff")
```